Study of the Detection System for Onscreen Contents Shooting (Countermeasure against Information Leakage by Video Recording/Photo Shooting)

Masaki Fujikawa, Hikaru Mori, and Kenji Terada

***Abstract*—With the expanded use of the smartphone and digital camera, unauthorized bringing out of information by photo/video shooting is reported. The authors have investigated whether detection of photo/video recording is possible focusing on the motion of a person and the feasibility of a system that quickly detects recording of the contents displayed on the screen. The system detects a camera held in one hand, both hands, and a fixing device, and uses algorisms to prevent erroneous detection and missed detection. In the experiments conducted for verification of the detection capability of the prototype system, highly accurate detection capability was demonstrated.**

***Index Terms*—Contents protection, information leakage, Security system, video/photo shooting.**

I. Introduction

With expanded use of the smartphone and digital camera it became quite easy to photo shoot static images or record movies, but on the other hand, cases of unauthorized bringing out of information by photo/video shooting are reported [1]. The majority of cases of bringing out information are made by the insider where the contents displayed on the PC screen (design drawing, customer information, etc.) are recorded as an image and brought out by uploading such image to online storage or recording it in portable memory.

The authors propose two countermeasures for photo shooting of the contents [2] and recommend using both countermeasures in a real application. The first one is a transparent infrared (IR) rays emitting sheet that is attached to the screen and the other one is a system that detects the photo/video shooting action.

The first one is used to make the contents recorded as static images or movies illegible utilizing the characteristic of the image sensor of the digital camera to recognize IR rays as ordinary light and introducing optical noise in the static image or movie of the recorded contents. A phenomenon where the human eye cannot recognize IR rays but the digital camera can is utilized.

The second countermeasure is used to make photo/video shooting impossible using an IR cutoff filter to circumvent the first countermeasure (i.e., a method to use an optical filter that selectively passes visible light while blocking IR rays in photo/video shooting the contents displayed on the screen to obtain static or video images that are not disturbed by optical noise), which detects the photo/video shooting activity from the motion of a person near the display screen.

The authors conducted a study of the feasibility of the second countermeasure, the details of the study are explained in this report, and the detection capability of the prototype system was evaluated by experiments. Discussions are made in the following order in this report. In Section II, differences of this experiment from other related research are explained, and then 1) conditions susceptible for bringing out information by photo/video shooting, 2) discussions on the motion of a person (camera holding), 3) premises for the system are explained. In Section III and IV, details of the system are explained. Section V provides an explanation of the experiments conducted for verification of accuracy in detection of the system and the results. Section VI provides a discussion for improvement of the prototype and for expanded use of the system.

II. Preparations

A. Difference from the Related Researches

Reference [3] is research with the same object as the authors. The concept to emit IR rays from the display screen to make the contents recorded illegible is the same. The authors' method is different in the approach to protect against photo/video shooting using an optical filter that blocks IR rays. In reference [3], the presence of the optical filter in front of the camera lens is detected by sensing retro-reflected IR rays at the display (an IR camera installed at the bottom of the display) utilizing the property of the optical filter to selectively mirror and reflect the IR rays (see Fig. 1).

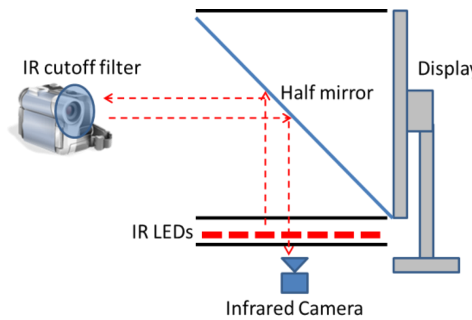


Fig. 1. Detection of a reflection type optical filter.

The above approach can detect the presence of the IR

Manuscript received September 6, 2013; revised March 7, 2014. This work was supported by the New Generation Information Security R&D Project of the Ministry of Economy, Trade and Industry (METI) of the Japanese Government.

M. Fujikawa is with the Sohgo Security Services Co., Ltd., Koto-ku, Tokyo 135-0014, Japan (e-mail: fujikawa-masaki@alsok.co.jp).

H. Mori was with the University of Tokushima, Tokushima Pref., 770-8506, Japan.

K. Terada is with the University of Tokushima, Tokushima Pref., 770-8506, Japan.

DOI: 10.7763/IJIEE.2014.V4.444

reflection type optical filter. However, an IR absorption type optical filter is also used (e.g., reference [4]), and as retro-reflection of IR rays from such a filter is very small, it is considered difficult to detect such a type of a filter by this method.

Furthermore, when a reflection-type optical filter is installed in front of the image sensor as in a single-lens reflex digital camera (see Fig. 2), as the area of the optical filter is very small, it is considered difficult for the IR camera in Fig. 1 to capture the retro-reflection of IR rays.

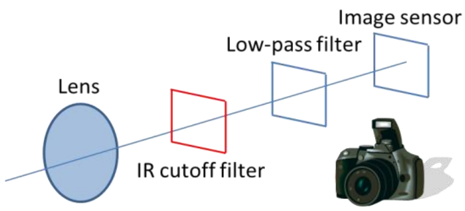


Fig. 2. Optical system of the single-lens reflex digital camera: These elements are incorporated in the camera.

Similar to reference [3], the authors investigated a method where detection of the optical filter and single-lens reflex digital camera adjacent to the display is possible when irradiated by IR rays from the display. (For example, there is a method where a camera installed on top of the display captures IR spectral images to detect the optical filter using the image processing technique [5].) However in this method, the lenses of glasses and the nameplate may be detected by mistake and will cause a problem in actual application.

Accordingly, the authors took a different approach and investigated whether a body activity for photo/video shooting could be detected focusing on human motion. The result indicated that any person who intends to photo/video shoot the contents displayed on the screen held the camera firmly using one hand or both hands or on a tripod to prevent the camera from moving.

Recently, a technology that allows noncontact and high-speed capture of human motion is available, accordingly if it is possible to digitally detect that a person is holding a camera, the system the authors are considering can be constructed. The relatively inexpensive motion capture system that detects motion of hands, feet, etc., as digital data without attaching markers to the human body (Kinect) is commercially available [6]. The authors decided to develop a prototype system using the Kinect to investigate the feasibility of the study where 3-D depth data captured by the Kinect are used to detect the motion of photo/video shooting (camera holding). As this approach is intended to detect human motion for photo/video shooting in general, it is considered that the method is also effective in the detection of photo/video shooting using a cutoff filter or single-lens reflex camera.

B. Discussion 1: Conditions Susceptible for Photo/Video Shooting of the Contents

In the office nowadays, the information accessed by the employee is recorded in an access log. Accordingly, it can be considered that bringing out of information by photo/video shooting of the contents onscreen will be attempted during office hours when such activity is unlikely to draw attention rather than at night or during holidays when such activity is easily noticed. In fact, many cases of information leakage by the employee occurred during office hours [7]. Photo/video shooting will be attempted so as not to draw attention or suspicion by other employees nearby pretending engagement in deskwork. (For example, photo/video shooting is made with the employee sitting at the desk rather than standing.)

Based on these assumptions, the authors considered that photo/video shooting is likely to occur when the employee is sitting at the desk for work and tried to detect the motion of the human body digitally under the condition of holding a camera.

C. Discussion 2: Motion for Photo/Video Shooting (Camera Holding)

When a person is holding a camera with one hand or both hands or using a tripod, hands or a tripod holding a camera will be naturally positioned in front of the body. This posture is due to one's unintended movement to confirm that the object is captured in the view finder (hole or LCD screen). When an employee is sitting at the desk for work, an elbow needs to be placed on the desktop or a small tripod needs to be used, and the above posture will become noticeable. When a person is at work with a PC, the arms will be naturally placed on the desktop in front of the body to operate a keyboard and a mouse. This means that when a certain object involving linear portion appears in the area in front of the body (i.e., the area close to the display screen) and such object remains stationary for a while, it can be decided that a camera is held using a forearm or tripod for photo/video shooting of the onscreen contents, and it is desirable to proactively protect the contents by turning off the display etc.

On the other hand, a person may sometimes rest an elbow on the desk during deskwork, and if contents are protected by mistake due to such motion of the forearm, work efficiency will deteriorate. The system should incorporate a function not to interpret such motion as photo/video shooting erroneously.

D. Premises

The following premises were used in this study.

- The range where holding of a camera can be detected is determined by the specification of the motion capture system. The Kinect was used as the motion capture and "Near mode" was employed to detect the motion of a person explained earlier, which can detect an object within approximately 40 cm to 3 m range from the Kinect.
- The skeletal detection application, which is the standard software of the Kinect, was not used. This is because the application cannot detect detailed information, such as a motion, shape, or irregularity of a body part, while it detects the gesture and pose of a human body in general. The point of this study is not to detect a gesture and pose of a human body.
- Holding of a camera with one hand, both hands, or by fixing device (tripod etc.) is considered in this study and other method of photo/video shooting (e.g. photo/video shooting by a micro camera attached on garment or glasses, by a camera with telephoto lens attached and placed in remote location from the desk using a remote controller, etc.) is not considered.

III. Development of Prototype System (System Configuration)

The prototype system was constructed according to the following procedure in this study in order to investigate the feasibility of the proposed system.

Fig. 3 shows the hardware that constitutes the prototype system. In the prototype, a display used for daily work by the employee is used as a component of the system. The Kinect connected to the PC is attached on top of the display. The reason the Kinect is attached at the top of the display is to avoid disturbing the operation of buttons located at the bottom of the display and to effectively utilize the detection range of the Kinect (approximately 57 degrees horizontally and 43 degrees vertically).

As shown in Fig. 4, the software was installed on the PC with the functions of capturing 3-D information from the Kinect, detecting camera holding, and controlling the software and hardware. The software, obtaining 3-D information from the Kinect, when it detects a camera is held, automatically turns off the display to protect the contents and notifies the administrator by e-mail that photo/video shooting of the contents is detected.

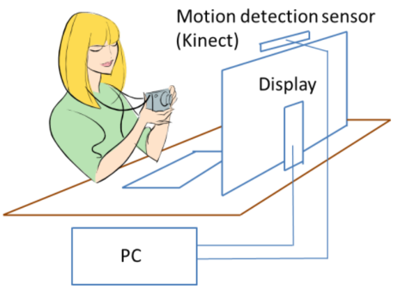


Fig. 3. Hardware configuration of the prototype system.

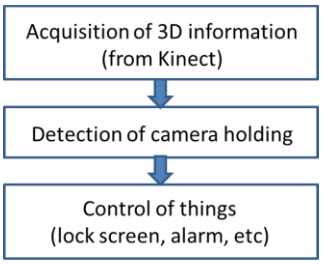


Fig. 4. Software configuration of the prototype system.

IV. Development of Prototype System (Software for Detection of Camera Holding)

In this section, an explanation of the software for detection of camera holding is given, which is the core of the system. This software is shown in the center box in Fig. 4 and consists of four processes as shown in Fig. 5.

The software upon acquiring 3-D information, evaluates the likelihood whether the 3-D object adjacent to the display is a forearm or a tripod holding a camera in position in Evaluation (1) focusing on the center of gravity, shape, and irregularity of the object. As in Evaluation (1), other body parts may be erroneously recognized as a forearm, re-evaluation is made in Evaluation (2), and at the same time, it is determined whether the motion of the object is not just resting an elbow on the desk similarly to holding a camera. In Evaluation (3), it is decided whether the 3-D object selected in Evaluation (2) is a forearm or a tripod holding a camera by measurement of the linear component included and residence time. A detailed explanation of the respective processes is given in the following sections.

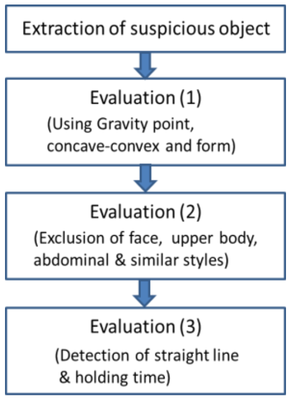


Fig. 5. Four processes that constitute the software.

A. *Extraction of Suspicious Object*

In this process, 3-D information continuously acquired from the Kinect is converted to the depth map sequentially, and the area for the 3-D object appeared near the display is extracted from the depth map assuming that it is a forearm or a tripod holding a camera. A depth map is an image where depth information (= distance information in the Z-axis direction) is converted into a 256 gray scale image as shown in photos to the right in Fig. 6 and Fig. 7. In the case of the depth map in Fig. 6, the area enclosing a cellphone camera is extracted as the 3-D object and in the case of the depth map in Fig. 7, a head is extracted as the 3-D object.

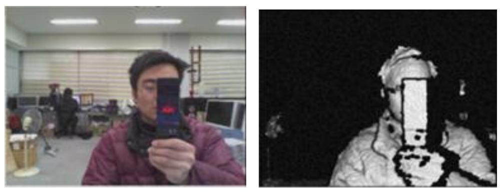


Fig. 6. A person holding cellphone camera Left: RGB image, Right: Depth map.

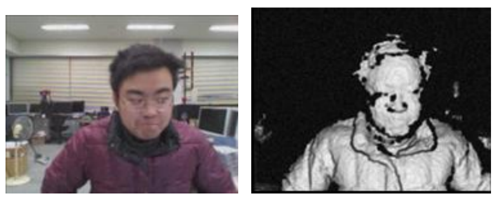


Fig. 7. A person bringing a head close to display. Left: RGB image, Right: Depth map.

B. *Evaluation (1)*

In this process, the likelihood as to whether the 3-D object extracted in the foregoing section (depth map information) is a forearm or a tripod holding a camera is evaluated by data processing in two steps.

1) *First step (Evaluation based on center of gravity)*

When a forearm section is divided in upper and lower halves with its center of gravity at the center, difference of the areas in both sections is small (it does not depend on inclination of the forearm, vertical, horizontal, or slanted). Accordingly, the difference between the two areas of the 3-D object (information of the depth map) divided into upper and lower sections at the calculated center of gravity should be investigated. If the characteristic explained above is observed, the determined likelihood that the extracted 3-D object (information of the depth map) is a forearm is high, and weighting should be applied.

Here, the y-coordinate of the center of gravity Y_g can be obtained by the following equation (the reason to obtain y-coordinate is to divide the object to upper and lower two sections). Where n is the number of pixels of the 3-D object (information from the depth map), and Y_i is the y-coordinate of each pixel.

$$Y_g = \frac{1}{n}\sum_{i=0}^{n-1} Y_i$$

2) Second step (Evaluation by shape and angle)

1) It is investigated whether the object can be divided into upper and lower sections by the section where depth information is not included as a boundary, focusing on the 3-D object (information of the depth map). An example of such an object is a head and body. Such a phenomenon occurs because IR rays from the Kinect are blocked by the presence of the jaw, and depth information at the throat becomes hard to obtain.
2) If the lower section is larger than the upper section with respect to the two sections divided as above (a) and the upper section is smaller than the threshold (i.e., two sections are the head and body), the determined likelihood that the sections are part of the forearm or a tripod is low, and no weighting is applied.
3) One side of the forearm or tripod is longer than the other side as in a rectangle. Because of this feature, whether the sections divided as in (a) above have such feature of a rectangle. If such a feature is observed, the likelihood that the section is a forearm or a tripod is high, and weighting is applied.
4) Irregularity of the surface of the forearm or tripod is smaller than the face, including parts that are not flat. Because of this feature, the distribution of angles in the sections divided as in (a) above is investigated. This is a method called shading, and the distribution can be obtained by calculating the angle formed by the normal vector $\vec{N}$ of the microplane forming the section (a plane formed by a pixel and the adjoining three pixels) and the illuminant vector $\vec{V}$ (IR ray emitted from the Kinect). If the above characteristic is observed, the likelihood that the object is a forearm or a tripod is high, and weighting is applied.

The x, y and z components of the normal vector $\vec{N}$ of the microplane formed by adjoining three pixels $(x1, y1, z1)$, $(x2, y2, z2)$ and $(x3, y3, z3)$ are expressed by the following equations.

$$N_x = (y2 - y1)(z3 - z2) - (z2 - z1)(y3 - y2)$$
$$N_y = (z2 - z1)(x3 - x2) - (x2 - x1)(z3 - z2)$$
$$N_z = (x2 - x1)(y3 - y2) - (y2 - y1)(x3 - x2)$$

The x, y and z components of the illuminant vector $\vec{V}$ can be considered 0, 0, 1. Then $\cos\theta$ of an angle formed by the normal vector $\vec{N}$ and illuminant vector $\vec{V}$ is obtained as follows.

$$\cos\theta = \frac{\vec{V}\cdot\vec{N}}{|\vec{N}|}$$

The color code corresponding to the angle calculated in (d) above is applied in Fig. 8 and Fig. 9, which indicates that an angle of 60 degrees or less is represented in gray, an angle of 60–80 degrees is represented in blue, and an angle of 80–90 degrees is represented in red.

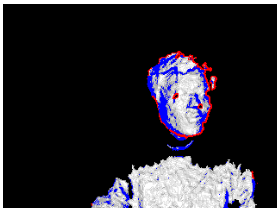


Fig. 8. Face and body. The likelihood that the object is a forearm or tripod is low according to (b), (c), and (d) of the second stage.

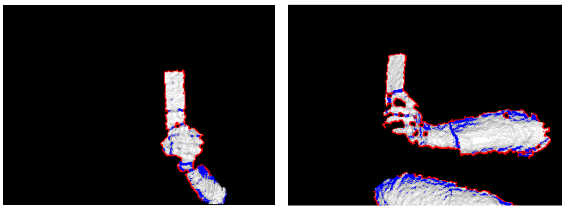


Fig. 9. Forearm holding a cellphone camera. The lower part of the photo to the right shows a left forearm. The likelihood that the object is a forearm or tripod is high according to (b), (c), and (d) of the first and second stages.

C. Evaluation (2)

As the posture of a person sitting at the desk is not constant and frequently changes, it is not uncommon that the 3-D object evaluated as a forearm or a tripod in Evaluation (1) is actuary other part of the body. According to the study made by the authors, it is possible that the upper half of the body, head, or abdomen excluding the forearm, or a posture resting the elbow on the desk, a similar position of the forearm holding a camera may be determined as a forearm or a tripod. Accordingly in Evaluation (2), the 3-D object is evaluated again and when the result indicates that the object is not a forearm or a tripod, processing is stopped and the system returns to acquisition process of 3-D information.

1) Case 1 (Upper half of the body excluding an arm)

Fig. 10 is a depth map of a person sitting at the desk with both arms behind the body, and pixels with depth information are colored in white for easy identification. Depth information at the portion of a throat is usually difficult to acquire, depth information may be available when a face is turned upward. In this case, a center of gravity is created at that point (red spot in the figure) according to (a) in Evaluation (1) and because difference in areas between the upper and lower parts is small, the image is evaluated as a forearm.

A reevaluation is made as follows. The maximum length in the vertical and horizontal directions (h and w respectively) and the number of pixels c that form a boundary between upper and lower sections should be obtained at the portion colored in white and if all *h*, *w* and *c* values are larger than the specified threshold values *h1*, *w1*, *c1*, it is decided as the upper half of the body excluding arms.

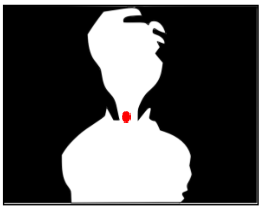


Fig. 10. Image of Case 1 and the center of gravity.

2) Case 2 (Head)

Fig. 11 is a depth image when a person brings the face towards the display and is colored in white similarly to Case (1). A center of gravity is created at that point (red spot in the figure) near the center of a head (large area in the figure) according to (a) in Evaluation (1). As the difference in area between the upper and lower sections of the center of gravity is small as in Case (1), this image is evaluated as a forearm.

A reevaluation is made as follows. With the area enclosed by a vertical rectangle, the long side of the rectangle *l* and short side *s*, and the ratio between white pixels and black pixels enclosed by the rectangle should be obtained, and if they are larger than the threshold values *l1*, *s1*, and *ratio1*, it is decided that the object is a head.

Fig. 11. Image of Case 2 and the center of gravity.

3) Case 3 (Abdomen)

Fig. 12 is a depth image when a person is sitting at the desk and bending backward on the seat back and is colored in white similarly to Case 1. In Evaluation (1), the center of gravity (red spot) comes near the center of the abdomen (large area in the figure). As the difference in area between upper and lower sections of the center of gravity is small like in Case (1), this image is evaluated as a forearm.

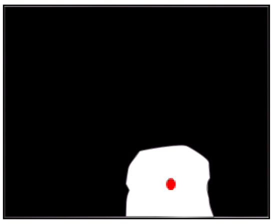


Fig. 12. Image of Case 3 and the center of gravity.

Reevaluation is made as follows. With the area enclosed by a horizontal rectangle, the long side of the rectangle *l* and short side *s*, and the ratio between white pixels and black pixels enclosed by the rectangle should be obtained, and if they are larger than the threshold values *l2*, *s2*, and *ratio2*, it is decided that the object is an abdomen.

4) Case 4 (Resting an elbow on the desk)

Fig. 13 is a depth image of a person resting the elbow on the desk, and is colored in white similarly to Case 1. In Evaluation (1), the center of gravity (red spot) comes near the center of the large area in the figure (area containing the head and forearm). Similar to other cases because the difference in area between upper and lower sections of the center of gravity is small, this image is evaluated as a forearm.

A reevaluation is made as follows. Maximum length *h* of the area in the vertical direction and the number of pixels with an angle 60 to 80 degrees *c* (i.e. parts constituting a face) are obtained, and when they are larger than the threshold values *h2* and *c2*, it is decided that a person is resting an elbow on the desk.

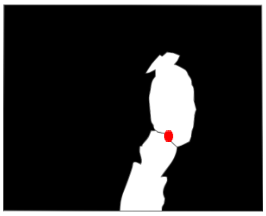


Fig. 13. Image of Case 4 and the center of gravity

D. Evaluation (3)

In this process, whether the linear component (i.e. the forearm or a tripod) is included and how long such linear component is detected are investigated for the 3-D object reevaluated in Evaluation (2) and determined not to correspond to Cases 1 through 4 to determine whether the object is a forearm or a tripod holding a camera. Investigation of presence of linear component is made by the commonly used Hough transformation. When the time interval during which a linear component is detected exceeds the threshold value *time*, it is determined that the likelihood of holding a camera with a forearm or a tripod is high.

V. EXPERIMENT

The purpose of this study is to investigate the feasibility as to whether camera holding can be detected. Accordingly in this section, whether camera holding using one hand, both hands, or fixing device can be detected by the prototype developed was verified by the experiment.

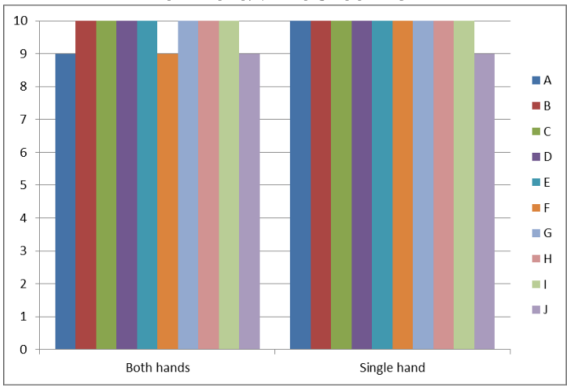
A. Experiment 1 (Camera Holding with one Hand or Both Hand)

First, to what extent camera holding with one hand or both hands can be detected was investigated. The authors gave an instruction to 10 subjects (A through J) to perform photo/video shooting 10 times using one hand and both hands, respectively, according to the following procedure.

> Sit at the desk. Then take out the photo/video shooting device (digital camera or smart-phone) that the subject uses daily from the desk drawer and hold it in a natural posture (not in unordinary posture), and photo/video shoot the contents (text data) displayed on the screen in front. Next, stow the photo/video shooting device in the desk drawer and leave the desk.

Table I shows a bar chart indicating number of successful detections for 10 times of photo/video shooting by the method of camera holding, and the left side chart indicates detection of camera holding by both hands, and the right side chart indicates detection of camera holding by one hand. The ordinate indicates the number of successful detections for 10 times of photo/video shooting, and the abscissa represents the respective subjects. The graph shows that the prototype system accurately detects camera holding with one hand or both hands.

TABLE I: BAR CHART INDICATING SUCCESSFUL DETECTION FOR 10 TIMES OF PHOTO/VIDEO SHOOTING



B. *Experiment 2 (Camera Holding with Fixing Device)*

Next, to what extent camera holding using a fixing device can be detected was investigated. In the experiment, a small tripod and a camera stand using a PET bottle were used as a fixing device (see Fig. 14).

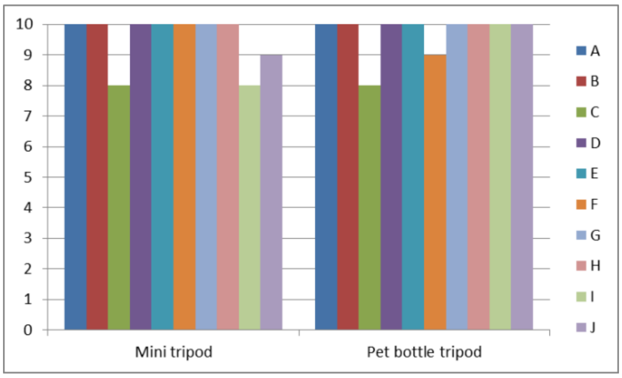
Fig. 14. Camera fixing device used for the experiment (Left: Small tripod, Right: Stand using a 500 ml PET bottle).

The authors gave an instruction to 10 subjects (A through J) to perform photo/video shooting 10 times using a small tripod and a PET bottle stand according to the following procedure.

Sit at the desk. Then take out the photo/video shooting device (a compact digital camera held by the fixing device) from the desk drawer and set it up on the desk, and photo/video shoot the contents (text data) displayed on the screen in front. Next, stow the photo/video shooting device in the desk drawer and leave the desk.

Table II shows a bar chart indicating the number of successful detections for 10 times of photo/video shooting for each fixing device, the left side chart indicates detection when a small tripod was used, and the right side chart indicates detection when a PET bottle stand was used. The ordinate indicates the number of successful detections for 10 times of photo/video shooting, and the abscissa represents the respective subjects. The graph shows the prototype system accurately detects camera holding using a holding device similarly in the case of Experiment 1.

TABLE II: BAR CHART INDICATING NUMBER OF SUCCESSFUL DETECTION FOR 10 TIMES OF PHOTO/VIDEO SHOOTING



VI. DISCUSSIONS

In this section, discussions are made to improve the feasibility of this study based on the results of experiments in section 4 and other related experiments.

A. *Reason of Detection Failure*

In Experiment 1 and Experiment 2, there are some cases where camera holding detection failed, and such failure is related to the limit of detecting range of the Kinect. Near Mode of the Kinect was used because the subject of this study attempted photo/video shooting during deskwork (i.e., a camera should be present near the display screen), and in all the cases where the detection of camera holding failed the distance from the Kinect to the camera was less than 40 cm and depth information of the camera and forearm was not adequately acquired (see Fig. 15).

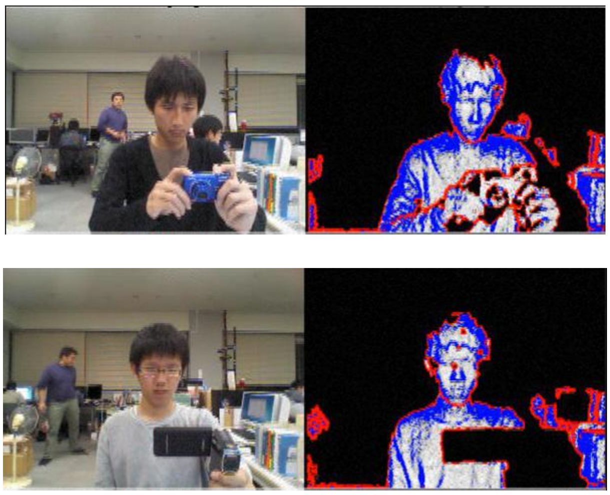


Fig. 15. A case where camera holding is detected (top) and a case where detection failed (bottom). As shown, depth information of a hand and a camera is missing because distance from the Kinect is too small.

The authors do not have a motion capture system that can detect a depth less than 40 cm this time, and it is required to repeat the experiments using such motion capture system to enhance the feasibility of this method.

B. Photo/Video Shooting Method to Avoid Detection and Countermeasures

When an employee realizes that the system can detect camera holding by one hand, both hands, or a tripod but cannot detect the posture of supporting an elbow on the desk, an employee may try to photo/video shoot the contents holding a camera with the arm in a similar position like supporting an elbow on the desk without confirming the object in the camera finder as shown in Fig. 16. According to the result of the investigation by the authors, 98% of the text data (text in two-text columns on one page of A4 size sheet) was legible when the contents were photo/video shot in this way.

Fig. 16. Photo/video shooting pretending that a person is simply resting an elbow on the desk.

The authors consider that a system to detect the presence of a camera is required to cope with photo/video shooting with a forearm holding a camera contacting the head. Specifically, installation of the following three algorisms is considered for the prototype.

- As the enclosure of the camera contains a flat surface, it should be investigated whether the 3-D object extracted (information of the depth map) contains angles with low values continuously. Distribution of angles can be calculated by shading explained in 4.2.2.
- As the enclosure of the camera contains a lens (a circular shape), it should be investigated whether the 3-D object extracted (information of the depth map) contains component of a circle. Component of a circle can be detected by Hough transformation.
- According to the study by the authors, while depth information of the enclosure can be obtained but depth information of the lens section is difficult to obtain for the single-lens reflex camera where the aperture of the lens is very large. Because of this phenomenon, it should be investigated whether the 3-D object extracted (information of the depth map) contains an area in which depth information is lacking.

C. Reduction of Erroneous Detection and Impact on Deskwork

The following experiments were conducted to investigate the motion of a body that induces erroneous detection and to investigate the impact of erroneous detection on the deskwork.

> The authors made 10 subjects who do not hold a camera engage in operations related to e-mail (for 10 minutes: drafting, editing, transmission, receiving, etc., of an e-mail) and in operation for software programming (20 minutes: generation, editing, debagging, etc. of a program). If erroneous detection occurs during the above operations, the motion of a body that caused erroneous detection was investigated, and the subjects were asked to what extent they were frustrated by erroneous detection.

The results are as follows:

Motion of a body that caused erroneous detection: Actions, such as making a prayer with both hands, pointing the display screen, or holding a head in both hands, are detected erroneously as the action of photo/video shooting. This is because linear component of the forearm was detected exceeding the threshold value *time*.

Degree of frustration: All the subjects responded that they felt frustrated, although the degree of frustration may be different by the phenomenon where information displayed on the screen disappeared all of a sudden (i.e., suspension of the operation). As frustration is a factor that decreases operating effectiveness, erroneous detection will affect deskwork.

At this moment, four cases described in 4.2.2 are excluded from the object for detection, but in the future, the motion of the forearm described above should be excluded from the object of detection by using the following three algorisms described in 6.2, and it is also considered effective to install the algorism in the prototype where the threshold value *time* can be changed flexibly depending on importance of the contents displayed on the screen.

D. Points to Expand Use of the System

The authors consider that the expanded use of the proposed system requires compact design, low cost, and enhancement of the capture function.

As BYOD (Bring Your Own Device) becomes popular, information leakage by photo/video shooting the contents displayed on the tablet device needs be considered and small and inexpensive motion capture device that can be installed in the tablet becomes necessary. If the low cost device can protect important information, it will help the expanded use of the system. As digital cameras today have a close-up mode, the system should preferably be capable of capturing a 3-D object about 10 cm from the surface of the screen.

The authors will continue to follow the trend of the motion capture device.

VII. Conclusion

The authors propose two methods as countermeasures for information leakage by photo/video shooting, and in this report, an explanation of the feasibility of the system to detect camera holding is given. According to the experiments using the prototype system, it was found that camera holding using one hand, both hands, or a tripod can be detected accurately.

The experiments also showed that suspension of the operation due to erroneous detection is a cause of frustration to the desk workers. Accordingly, improvement of the system needs to be considered to reduce impact on the deskwork due to erroneous detection by linkage with the algorism detecting the presence of a camera and by flexibly changing threshold value *time* depending on the significance of the contents.

References

[1] T. Kaiser. (2011). Air Force One Flight Plan Blogged by Tokyo Air Traffic Controller. *Daily TECH.* [Online]. Available:

http://www.dailytech.com/Air+Force+One+Flight+Plan+Blogged+by+Tokyo+Air+Traffic+Controller/article22691.htm

[2] M. Fujikawa, R. Kamai, F. Oda, K. Moriyasu, S. Fuchi, T. Takeda, H. Mori, and K. Terada, "Development of Countermeasure Systems for Content Leaks by Video Recording/Camera Shooting," in *Proc. the International Conference on Information Society* (*i-Society 2012*).

[3] National Institute of Informatics. Technology to prevent unauthorized copying of displays by utilizing differences in sensitivity between human beings and devices - Prevent disclosure of confidential and personal information through unauthorized copying of displays. [Online]. Available: http://www.nii.ac.jp/en/news/2011/0704/

[4] Sumita Optical Glass, Inc. Near Infrared Absorption Filter. [Online]. Available: http://www.sumita-opt.co.jp/en/functional/functional05.htm

[5] M. Fujikawa, J. Akimoto, F. Oda, K. Moriyasu, S. Fuchi, and Y. Takeda, "Study of Countermeasures for Content Leaks by Video recording," in *Proc. the 6th International Conference on Availability, Reliability and Security (ARES 2011)*, 2011.

[6] Microsoft. KINECT for windows. [Online]. Available: http://www.microsoft.com/en-us/kinectforwindows/

[7] Japan's black market for personal info…how safe are you? [Online]. Available: http://forum.gaijinpot.com/showthread.php?121846-Japan-s-black-market-for-personal-info-how-safe-are-you

Masaki Fujikawa was born on 6 February 1974 in Tokushima Prefecture, Japan. He received his master's degree in information engineering in 1998 from Tokushima University, Tokushima Prefecture, Japan. In 2004, he received his Ph.D. in information engineering from Chuo University, Tokyo, Japan.

He joined the Sohgo Security Services Co., Ltd, Tokyo Japan in 1998. From October 2002 to March 2011, he doubled as a guest researcher of the Research and Development Initiative, Chuo University. From November 2010 to Mach 2012, he was the director of an information security R&D project of METI (Ministry of Economy, Trade and Industry, Japan). Now, he is a senior researcher of the company, and researching a variety of security and safety systems in order to build a safer society.

Dr. Fujikawa is a member of IEEE, IPSJ and the steering committee of the IPSJ Computer Security Group. Dr. Fujikawa received the best paper award of the Computer Security Symposium 2013 and the best demonstration award of the 4th IFIP WG 11.11 International Conference on Trust Management 2010.

Hikaru Mori was born on 10 November 1988 in Tokushima Prefecture, Japan. He received his master's degree in information engineering from Tokushima University, Tokushima Prefecture, Japan. He joined the Terasaki Electronic Co.,Ltd, Osaka Japan in 2013. Now He is a switchboard designer of the company, and designing switchboard in order to build a safer society.

He received the best demonstration award in 2011 on Computer Security Symposium (CSS) 2011.

Kenji Terada was born in 1967. He received the B.E. degree from Keio University, Japan in 1995 and the Ph.D. degree in electrical engineering from Keio University.

He engaged in research in the University of Tokushima in 1995. He is currently a professor in the Faculty of Engineering, the University of Tokushima. Dr. Terada's principle research interests are image processing and image sensing. Dr. Terada is a member of IEICE, IEEE, SICE, IEEJ, IIEEJ, JSPE and ISCIE.